Understanding Emotions in SNS Images From Posters' Perspectives

Junho Song
Hanyang University
Seoul, South Korea
hahadori@hanyang.ac.kr

Kyungsik Han
Ajou University
Suwon, South Korea
kyungsikhan@ajou.ac.kr

Dongwon Lee
Pennsylvania State University
Pennsylvania, USA
dlee@ist.psu.edu

Sang-Wook Kim*
Hanyang University
Seoul, South Korea
wook@hanyang.ac.kr

ABSTRACT

As the popularity of media-based social networking services (SNS), such as Instagram and Snapchat, has increased significantly, a growing body of research has analyzed SNS images in relation to emotional analysis and classification model development. However, these prior studies were based on relatively small amounts of data, where the emotions of images were labeled from viewers' perspectives, not posters' perspectives. Consequently, we analyze 120K images that reflect poster's emotion. We develop color- and content-based classification models by considering: (1) the dynamics of SNS, in terms of the volume and variety of images shared, and (2) the fact that people express their emotions through colors and objects. We demonstrate the comparable performance of our model with models proposed in prior studies and discuss the applications.

CCS CONCEPTS

• **Applied computing** → Arts and humanities; • **Information systems** → Sentiment analysis;

KEYWORDS

Image-based emotion analysis, classification, social network service

ACM Reference Format:
Junho Song, Kyungsik Han, Dongwon Lee, and Sang-Wook Kim. 2020. Understanding Emotions in SNS Images From Posters' Perspectives. In *Proceedings of ACM SAC conference (SAC'20).* ACM, New York, NY, USA, Article 4, 8 pages. https://doi.org/10.1145/3341105.3373923

1 INTRODUCTION

Social networking services (SNS) have become an important part of people's lives in terms of forming and maintaining social relationships [16, 34, 35]. SNS have greatly impacted society (e.g., politics, economy), beyond relationships based on individual social activities among users. As a result, SNS influences are being used for publicity and marketing purposes [11, 25]. Users share information about themselves or activities and express their thoughts, feelings, and opinions in an explicit or implicit fashion.

In addition, the form of sharing has been changed from typical text-based to media-based communications [15, 24, 29]. The recent surged popularity of Instagram or Snapchat, an image-based SNS, demonstrates this trend. In this respect, the attempt to grasp human characteristics through image-based SNS, in line with recent trends, is being actively made in the multimedia community.

Emotion is an important part in understanding human characteristics, because it affects human behavior and thinking [10, 13]. Many prior studies have suggested that various features (e.g., colors, contours, textures) can be extracted from images and presented in emotion classification models [3, 5, 7, 21–23, 32, 36–38]. The literature review, however, revealed two important aspects that have been somewhat neglected for a better analysis of SNS images.

The first aspect is that existing research is based on a small amount of data. In addition, the nature of these datasets is different from that of SNS images. Hence, these small datasets are not likely to reflect the characteristics of SNS, regarding the volume and diversity of the images therein. In fact, the numbers of images from the IAPS [19], ArtPhoto [23], and Abstract Painting [23] datasets, which were widely used in prior research as a benchmark dataset, are 395, 806, and 228, respectively. Furthermore, the emotions in those images could easily be recognized.

The second aspect is that since the numbers and types of images shared by SNS are very large, the emotions can also be expressed through various images. To construct an emotion classification model, images labeled with ground truth emotions are required. For this purpose, labeling has been performed in accordance with the emotion of a third person who views the image. However, the potential problem with this method is that the data could only be composed of images with emotions that are somewhat easy for viewers to determine. For example, from a third party's perspective, the image of people at the park in daytime may suggest feelings of happiness and pleasure for viewers. However, a poster who uploaded the image may have wanted to express sadness and loneliness paradoxically (the image may contain an empty bench that implies the emotion). For this reason, existing emotion classification models may have limited functionality for real, various SNS

*Corresponding author

SAC '20, March 30-April 3, 2020, Brno, Czech Republic

ACM ISBN 978-1-4503-6866-7/20/03…$15.00
https://doi.org/10.1145/3341105.3373923

images. Thus, it would be necessary to build a model focusing on the emotions of image posters, in addition to image viewers.

Our work focuses on the aforementioned limitations. More specifically, we use a large-scale dataset (i.e., 120,000 images) that reflects image posters' emotions. We investigate six emotions (i.e., happy, sad, angry, scared, excited, satisfied) from the collected images, with respect to their colors and content. Literature illustrates that colors and content are closely related to the emotion of the person expressed in the images [9, 14, 32]. The information is used to develop emotion classification models. Because few studies have investigated the analysis and model development based on the data reflecting upon the content and type of the image object, our work presents new insights on emotions in SNS images. Consequently, we aim to answer the following two research questions.

- RQ1. Do the color and content of SNS images relate to the poster's emotions?
- RQ2. Can we create a classification model that predicts the poster's emotions based on the color and content in SNS images?

To answer RQ1, we analyze the color distribution of the SNS image corresponding to six emotions in RGB and HSV color space. The results reveal a difference in the color distribution corresponding to each emotion. The distribution in the HSV color space is significantly different from that in the RGB color space. In addition, the HSV color model is found to be better for the emotion classification. Through Latent Dirichlet Allocation (LDA) [4] topic modeling, we identified 15 topics from the image objects (i.e., inferring image content) and found significant differences in the topic distributions among six emotions.

To answer RQ2, we build classification models using the Neural Network, SVM, and Decision Tree. For the architecture of the neural network model [12], we merged three individual models (i.e., RGB-, HSV-, and object-based models) into one, leading to a performance increase. The overall model performance of the merged model was comparable to the most recently reported CNN-based emotion classification models.

Finally, we demonstrate the application of our model over a new test dataset. We illustrate that it would be difficult to identify the emotions in images through the viewers' perspectives and highlight the ability of our model for correctly detecting the emotions.

Our research contributes to an in-depth and improved understanding of the relationship between images and posters' emotions in the context of SNS. We highlight the importance in considering the image posters' emotions, which will better reflect the dynamics of SNS images in terms of volume and diversity, compared to the emotions from a third-party's perspectives. By taking this into consideration, it is expected that the classification models would be better and more suitably applied to a wider variety of fields. These classification models can help in the prediction of psychological change (e.g., depression) or the generation of a recommendation system (e.g., advertisements, music, games), where the consideration of a user's own emotions as a poster is important.

2 RELATED WORK

Emotion is the generic term for the subjective, conscious experience that is primarily characterized by psychological expressions, biological reactions, and mental states [13]. The research that relates to emotions from images has been conducted for a long time. Since the sharing of images through SNS is highly active, the importance of emotional analysis through images is increasing.

2.1 Correlating emotions with image features

Colors are the most common way to identify photo features. Patricia and Mehrabian studied the relationship between emotions and the features of colors [33], including the brightness of achromatic colors and hue. They found a stronger effect of brightness than saturation in determining the dominance response to color. Barry, Hardesty, and Suter studied the relationship between consumers' reactions, shopping intentions, and various combinations of color and light. Orange and blue were studied for colors, and bright and soft were studied for lights [2].

NAz and Epps showed that the relationship between color and emotions is closely tied to color preferences [26]. They also found a contextual aspect (e.g., culture, age, racial group) of color preferences that could be interpreted in a different fashion. Machajdik and Hanbury defined a comprehensive set of colors and color-related features for the relationship with emotions including color types (e.g., black, blue, brown, grey, orange) such as, brightness, saturation, hue, colorfulness, dark, light, warm, and cold [23].

Many studies have also utilized the various characteristics of image object information. Machajdik and Hanbury analyzed human-related features, including faces and skin [23]. Face features include the number of frontal faces, relative size of the biggest face, etc. Skin features include the number of skin pixels, relative amount of skin with respect to the size of faces, etc. Zhao et al. extracted image features using the principles-of-art-based emotion features (PAEF) [38]. PAEF is divided into several categories: balance, emphasis, harmony, variety, gradation, and movement. Of these, balance and movement were foud to vectorize the object's composition and movement degree in the picture. Lu et al. identified the shapes of the objects and associated them with emotions [22]. The shape features included line segments, angles, and continuous lines. The results show that angry is related to the circularity of the image, disgust is related to the length of the line segments, and both awe and excitement are related to the orientation of the line segments.

2.2 Developing emotion models

A number of studies have proposed models that classify sentiments (positive, negative) or emotions (types of emotions vary by study) from images.

For emotion classifications, many studies have constructed a benchmark comparing their models with others using three datasets: IAPS (394 images), ArtPhoto (806 images), Abstract (228 images). These datasets all include 8 emotions: amusement, anger, awe, contentment, disgust, excitement, fear, and sad. Other studies classify these 8 emotions into two sentiment classes (i.e., positive and negative) and evaluate their models. The evaluation metrics vary by study, where accuracy, average true positive rate, or area under the curve (AUC) have been used.

Many studies have employed traditional machine learning models for classifying sentiments or emotions. Machajdik and Hanbury developed methods to extract and combine features that represent the emotional content of an image (e.g., color, texture, composition, content) [23]. Their models are based on the Naive Bayes.

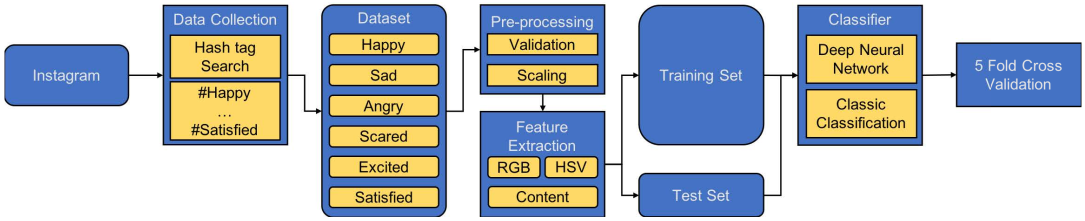


Figure 1: Study Procedure.

The model's performance was better than that of the SVM or the Random Forest, with an average true positive rate between 60-70% for each emotion.

Lu et al. focused on the relationship between shapes and emotions [22]. Regarding shapes, the authors leveraged psychological studies and found that the roundness and the complexity of shapes were fundamental to understanding the emotions. They used IAPS datasets, where three basic dimensions of affects (i.e., valence, arousal, and dominance) are specified. For modeling, however, they considered valence and arousal for the binary classification, which accomplished 76% accuracy. Similarly, Lu et al. established three new computational constructs (i.e., roundness, angularity, and complexity) and models that classify the valence and arousal of the images [21]. Their experimental results show that the color features related most strongly to the positivity of the perceived emotions. The texture features related more to calmness or excitement, while and roundness, angularity, and simplicity related to both of these emotional dimensions.

Borth et al. presented a method built on a psychological theory (i.e., Plutchik's Wheel of Emotions) and web mining to discover the sentiment words [6]. They constructed 3,000 ANPs (Adjective Noun Pairs), including beautiful flower, disgusting food, etc. Using the ANP detector, they built an application, called Sentibank that provides sentiment prediction in image tweets. The SVM model for classifying positive or negative sentiment, based on SentiBank and text features, yielded 72% accuracy. Zhao et al. studied the concept of the principles-of-art emotion features (PAEF) (e.g., emphasis, harmony, variety, gradation, movement), and its influence on image emotions. The types of emotions they studied include amusement, awe, contentment, excitement, anger, disgust, fear, and sadness [38]. By using SVM, they suggested a classification model with the average true positive rate between 60-70% for each emotion.

Recent modeling work includes the utilization of deep learning techniques, such as Convolutional Neural Network (CNN). For example, Xu et al. illustrated that the image representations from the CNN could be effectively transferred to a sentiment analysis [36]. They constructed a sentiment benchmark (positive and negative) and showed that the proposed model classified sentiments with the AUC of 64%. Chen et al. presented a multi-scale CNN, where the features were first extracted from the pre-trained ImageNet dataset, aggregated using the Fisher Vector for multiple levels, and then concatenated to form a compact representation [7]. They used ArtPhoto (conventional dataset) and FlickrEmotion (newly collected). The average true positive rates of the proposed model for each of the 8 emotions ranged from 70-82%. You et al. significantly expanded the size of the data by collecting a large number of weakly labeled emotion related images and asking crowd-workers to manually label the images [37]. They developed a fine-tuned CNN and compared its performance with other state-of-the-art models, which yielded 58.3% accuracy for classifying eight emotions (e.g, amusement, anger, awe, contentment, disgust, excitement, fear, and sadness). However, the number of samples per emotion was significantly unbalanced. Thus, the accuracy is expected to be lower than what was reported.

2.3 Research directions and goals

Based on our literature review, we developed many points that have yet to be considered, despite their importance, in the analysis of emotions in the context of SNS. We consider this as another research direction for better understanding the relationship between image features and emotions.

2.3.1 Image feature selections. The emotional relationship of people to things has been the same significance as the relationships between individuals and groups of people [9, 14, 32]. The research has illustrated that emotions are attached to objects, which also depict the emotional status of people and embody memories, providing a tangible link to the past. This applies to people who possess the objects, as well as those who look at the objects. Thus, in the context of social media, objects embedded in the images can provide a connection to people who post the content and to those who look at the content.

In research investigating the image object, model construction through the utilization of the shape and the color of the object has progressed. However, our literature review indicates that there is no model construction using the kind and meaning of the object. Based on this, an understanding of the relationship between the image posters' emotions, based on the combination of image object types and colors, is a research direction that yet to be implemented. This is the primary goal of our work.

2.3.2 Emotions of image posters. When it comes to emotional labeling, most studies have been conducted in such a way that the third parties evaluated the emotions of the image poster. While this is a valid approach, it also has the disadvantage of being able to concentrate only on the images, where the emotions that can be easily labeled by the third party. From the perspective of the image poster, the image itself may be different from the emotions that the poster has, which may be difficult, or not explicit enough for third parties to understand. For example, a person may post a bright photo with flower, where it is likely that the third party would think the image is related to positive emotions such as happiness. However, the photo poster may have feelings of sadness, which is difficult to grasp from the third party's standpoint.

The images shared in the SNS are very large, in terms of diversity and volume, which can be expressed through a variety of photographs to express emotions. Focusing on the images with explicit emotions does not seem to be comprehensive enough, in terms of understanding emotions from images in the context of SNS. Therefore, we collected images on Instagram by considering this and constructed the emotional models accordingly.

3 STUDY PROCEDURE

3.1 Data collection

We collected public photos of users in the photo-based SNS service, Instagram, according to the emotional category of previous research [10]. The collection was conducted between January and February 2018. The emotional category was divided into six main emotions: happy, sad, angry, scared, excited, and satisfied. This emotion group is widely applied in the field of psychotherapy. Emotional categories are employed in existing emotion model studies. We searched for photos on Instagram, based on the hashtags of each emotion. Here, prior studies have that the users have used the hashtags to express their intentions [1, 20, 27]. Several verification steps were conducted with the images: (1) we removed the images that contain conflicting tags (e.g., having #happy and #sad tags together); (2) we manually removed the spam-like photos (e.g., advertisements). From around 40,000 verified images per emotion, we randomly sampled 20,000 images, having a total of 120,000 images for the analysis.

3.2 Feature extraction

We extracted color and object information from the collected images. As previously mentioned, research has proven the influence of colors on user's intentions and emotions. Consequently, we used two color-based models: the RGB (Red, Green, Blue) model and the HSV (Hue, Saturation, Brightness) model. This is due to the intuition that brightness and saturation information will affect emotions. The RGB model is based on the Web Safe Color Table, which illustrates a total of 216 features that can be recognized by humans on the monitor. The HSV model uses total of 560 features; hue (360), saturation (100), and brightness (100) values.

Figure 2: Example of the result of image object extraction.

Our literature review indicates that although some of the object information (e.g., object colors, textual, shape, contour) were used in prior studies, the meaning of the objects has not been utilized as a feature for the emotion classification. Given that the objects in a poster's image are often related to the poster's intentions and emotions [9, 14, 32], it is necessary to create an emotion model based on this. Hence, to extract the object information from the image, we used the MS Azure Cognitive Service API [1]. The result of object extraction through the API is shown in Figure 2, each object has a confidence value. We only included objects with a confidence value over 70%. Overall, about 1,000 objects were extracted.

[1]https://azure.microsoft.com/en-us/services/cognitive-services/

3.3 Emotion model

We created emotion models based on the color and object features extracted from the images. We constructed various emotion classification models by using the classic classification algorithms (e.g., SVM, Decision Tree) as well as a neural network model, applying deep learning.

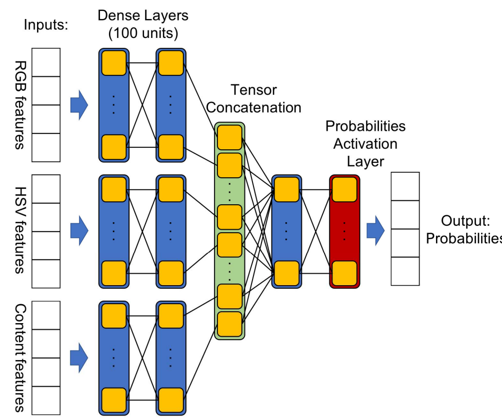


Figure 3: Neural network architecture for emotion classification used with colors and image content.

We experimented with several baseline models and developed the neural network architecture presented in Figure 3 using the Keras framework [2]. We constructed three separate networks for each primary features (i.e., color RGB, color HSV, image objects). The numbers of features for color RGB, color HSV, and image objects are 216, 560, and 1,082, respectively. We used the PCA [31] to reduce the dimensions of each feature to 50, here the number of dimensions was decided experimentally. We relied on the late fusion approach, that has been shown to be effective in many machine learning tasks [17]. "Fusion" allows for a network to learn a combined representation of multiple input streams. In our case, this fusion was executed in the classification layers. To combine these networks into one prediction and train them together, we merged these dense layers before the final classification. We trained our models for 100 epochs using the ADAM optimization algorithm [18] and evaluated them using 5-fold cross-validation. For traditional classification algorithms, we used the SVM and Decision Tree models. Each model uses 80% of the dataset as learning data, 20% as test data, and includes 5-fold cross-validation.

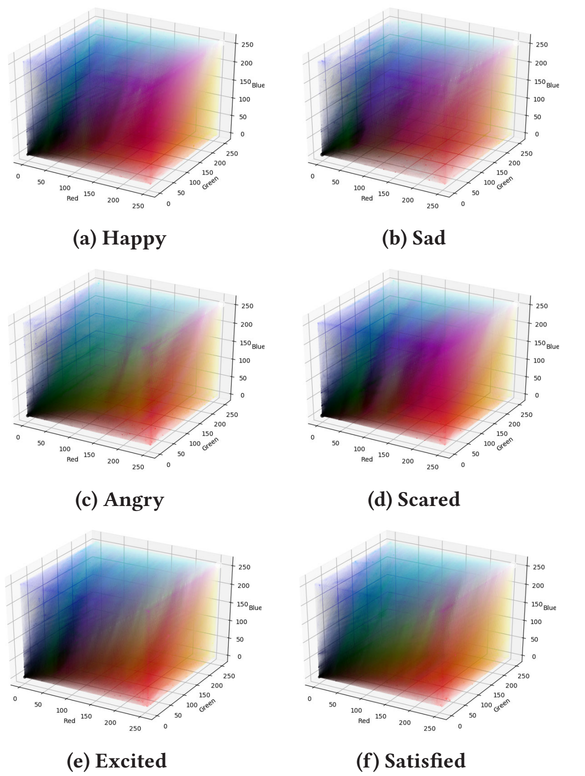
4 RESULTS

4.1 Colors of SNS images

To examine the relationship between the colors and the emotions in the SNS images, we extracted the RGB and HSV colors from all the pixels of the image and measured the frequency of each color.

Figure 4 illustrates the results of visualizing the color distribution and frequency of each emotion image in the RGB color space, which considers the specific gravity of each red, green, and blue). For all emotional images, the color frequency is high in black (about 9%) and white (about 10%). While prior studies presented similar color distributions among positive emotions (i.e., happy, excited, satisfied) and negative emotions (e.g., sad, angry, scared), our results

[2]https://keras.io/

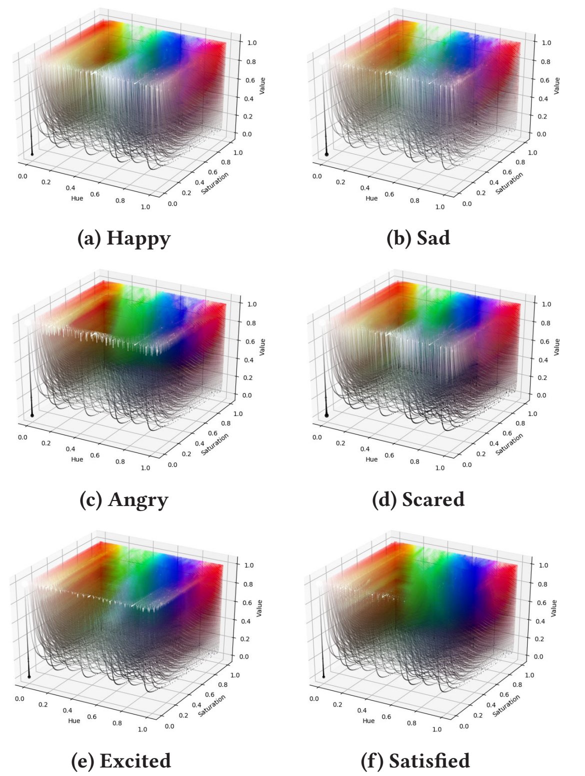


(a) Happy (b) Sad

(c) Angry (d) Scared

(e) Excited (f) Satisfied

Figure 4: Color distribution on the RGB color space. A group of [happy, sad, scared] and of [angry, excited, satisfied] show similar color distributions.



(a) Happy (b) Sad

(c) Angry (d) Scared

(e) Excited (f) Satisfied

Figure 5: Color distribution on the HSV color space. A group of [happy, sad, scared] and of [angry, excited, satisfied] show similar color distributions.

illustrate a similar distribution among [happy, sad and scared] and [angry, excited, and satisfied]. This results from the difference in the labeling method between the datasets (i.e., IAPS, ArtPhoto, and Abstract Painting) used in the previous research and the SNS image dataset used in our research. The emotion in the previous research is based on the viewer's standpoint, whereas that of the SNS image is based on the poster's intentions and messages through tagging.

We compared the color extracted from each emotion in the HSV color space, which considers the weight of each hue, saturation, and value. This was conducted to determine how the saturation and brightness differed from each emotion. As shown in Figure 5, similar to the results from the RGB colors, [happy, sad, and scared] and [angry, excited, and satisfied] showed similar color distributions. In the HSV, the colors were widely distributed, due to the differences in the saturation and the lightness. Clearly, the black color and the white color occupy a large portion of the entire color space. This is because the achromatic color produced by the difference in brightness between white and black is expressed by the characteristics of the color space that can express the brightness and color saturation.

4.2 Content in SNS images

Figure 6 shows the distribution of the top-20 content (based on frequency) across emotions. Because distributions vary by emotion, it can be inferred that the content of the image is related to the emotion. Here, the content with the highest weight for each emotion includes person, indoor, and outdoor. This may be due to the characteristics of SNS, since SNS users share their life activities. In

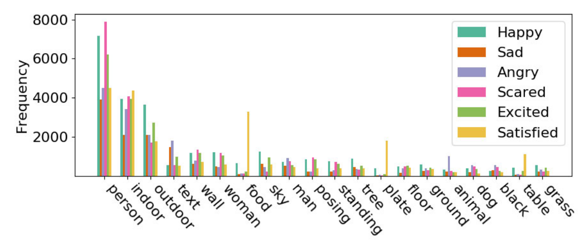


Figure 6: Distributions of top-20 image content.

addition, the users express happy and scared feelings through the person in the image, and satisfied expressions through food related images (e.g., food, plate, table).

Table 1: Content of topics identified through LDA.

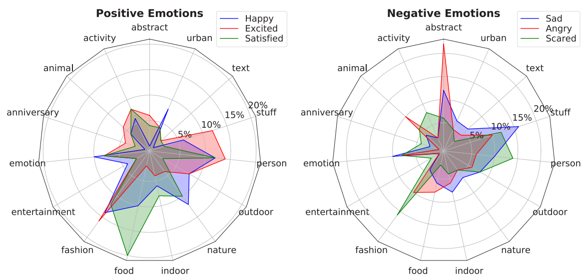
Topic	Content				
abstract	different	stationary	old	finned	hay
activity	travel	sport	pool	fishing	spectacles
animal	reptile	pet	cat	lion	dog
anniversary	birthday	decorated	party	ceremony	wedding
emotion	gloomy	red	bad	pink	green
entertainment	book	stage	screen	music	game
fashion	shoes	cloth	costume	glasses	accessory
food	fruit	bowl	drink	plate	porridge
indoor	kitchen	toilet	bathroom	window	floor
nature	valley	water	sky	grass	canyon
outdoor	road	park	bench	ground	tree
person	man	posing	prey	head	sitting
stuff	toy	oddments	clock	gift	goods
text	newspaper	sign	map	board	recipe
urban	city	car	building	street	transport

We examined the relationship between the content of each image and each emotion in more detail through topic modeling. We

Table 2: F1 score of emotion classification (DT: Decision Tree, NN: Neural Network).

	Happy			Sad			Angry			Scared			Excited			Satisfied		
	SVM	DT	NN	SVM	DT	NN	SVM	DT	NN	SVM	DT	NN	SVM	DT	NN	SVM	DT	NN
RGB	.68	.76	.77	.54	.79	.77	.75	.75	.76	.76	.78	.76	.75	.76	**.77**	.75	.80	.77
HSV	.76	.76	.77	.76	**.80**	.77	.75	.76	.76	.77	.78	.76	.76	.76	**.77**	.77	.80	.77
Content	.77	.76	.77	.76	.76	.77	.77	.77	.77	.77	.76	.76	.76	.76	**.77**	.83	.83	.83
Combined (RGB+HSV+Content)	.76	.77	**.82**	.76	**.80**	**.80**	.75	.75	**.79**	.76	**.79**	.76	.76	.76	**.77**	.77	.84	**.85**

used LDA [4]. We extracted 50 topic objects, grouped similar topic objects, and classified the objects into 15 topic objects. We repeated this process multiple times, until we reached the consensus set of topics. Figure 7 illustrates the results. Positive emotions and negative emotions showed different topic distributions. Positive emotions illustrated high topic distributions in the order of food, fashion, person, and activity, while negative emotions showed high topic distributions in the order of abstract, stuff, and emotions.



(a) Positive emotions **(b) Negative emotions**

Figure 7: Topic modeling results of positive and negative emotions.

We note that an animal topic provides a similar weight between positive emotions (8%) and negative emotions (10%). Animals with positive emotions include cats, dogs, parrots, and other pets. Those with negative emotions include lions, reptiles. This supports the claim that the content of the image represents the user's emotions.

4.3 Performance of the classification model

Table 3: Accuracy of emotion classification models.

Models	Features	Accuracy
SVM	Combined (RGB + HSV + Content)	30%
Decision Tree	Combined (RGB + HSV + Content)	33%
InceptionV3 (transfer learning)	Raw images	30%
VGG16 (transfer learning)	Raw images	31%
Xception (transfer learning)	Raw images	35%
Deep Neural Network	Combined (RGB + HSV + Content)	**39%**

To examine the performance of our emotion classification model using the colors and content of the images, we benchmarked our model with VGG16 [28], InceptionV3 [30], and Xception [8] models. These models are pre-trained on the ImageNet dataset. They are provided in the keras library for use. We applied a transfer learning technique, which freezes all but the penultimate layer. It re-trains the last dense layer to the model and obtains the results. The goal of transfer learning is to apply the knowledge learned from one environment to another environment. As shown in Table 3, the results indicate an increased performance of our models, compared to other models.

Table 2 summarizes the F1 scores of the models. Here, we show that the models based on combined feature (RGB+HSV+Contents) shows better performance than models based on single feature. Especially, the merged neural network model yielded better performance over all emotions than others. The highest improvement of performance is *sad* which is 22% (0.54 to 0.76) on SVM model. In addition, the neural network model shows a comparative performance to previous CNN-based emotional classification models using fine-tuned images (70-90% the average true positive rate, described in Related Work). Hence, we confirmed that the features that we analyzed are significant for emotion classification model.

4.4 Feature importance

We examined the importance of features used in the models.

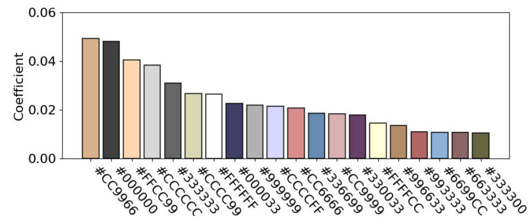


Figure 8: Feature importance of DT with RGB feature.

The most dominant colors in the RGB color space are the red colors (#cc9966, #ffcc99, #cccc99, #cc6666, #cc9999, #330033, #ffffcc, #996633, #993333, #663333), followed by black, white, and gray colors (#000000, #cccccc, #333333, #ffffff, and #999999) and the blue colors (#000033, #ccccff, #336699, and #6699cc) (Figure 8).

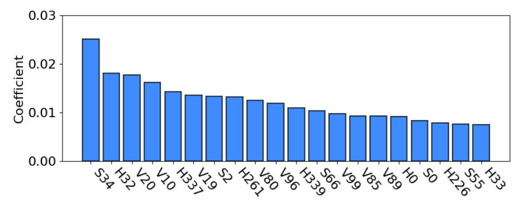


Figure 9: Feature importance of DT with HSV feature (H:Hue, S:Saturation, V:Value).

In the HSV color space, the important results are presented in Figure 9. The most dominant feature of the significance result is Value in HSV (v20, v10, v19, v80, v96, v99, v85, v89), because the sum of the coefficients was 0.0999. The sum of the coefficients of hue and saturation components were 0.0735 and 0.0648, respectively. This is the result of the saturation and brightness being related to the emotion, as mentioned in the feature extraction.

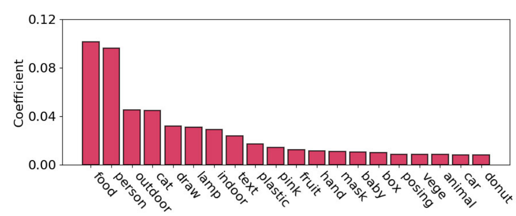


Figure 10: Feature importance of DT with content feature.

The most dominant characteristic of the contents (Figure 10) was food (0.101) and person (0.095). Other important features included fruits, vegetables related to food, and hands, babies, posing related to a person. In addition, outdoors, indoors, texts, cats (i.e., animals), stuff (i.e., lamp, plastic, mask), and colors (i.e., pink) were also important. Overall, topic modeling almost coincided with the topic that represents each emotion.

4.5 Model testing

Finally, we tested our model further on a new dataset of 600 samples labeled with the corresponding emotion. We examined the characteristics of the emotional images.

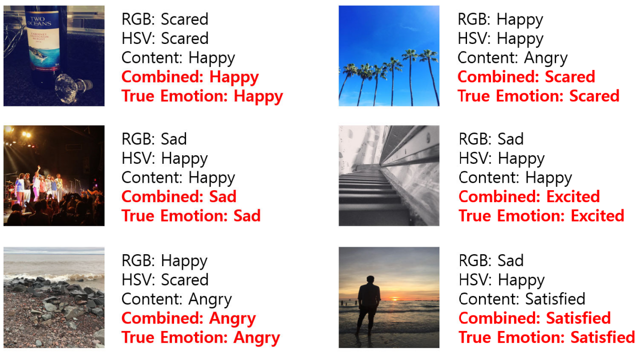


Figure 11: Examples of images for each of the six emotions

Figure 11 illustrates a correctly classified image for each emotion. We wanted to pay attention to the features of each image. When viewing an image, it was not clear how to derive the corresponding emotion for each photo. In this case, it was difficult to grasp the actual feelings of the poster. Since the type and number of images shared on the SNS is significantly large, the emotions of those who follows these images can also appear from image that we do not generally think of.

For example, consider the image with the emotion labeled in happy in Figure 11. Our model, based on the RGB and HSV color features, recognized the image as scared. Presumably, this result may be derived from dark colors in the image. Conversely, the results of the content and the combined features were happy. We could assume that the poster may take a rest, having a glass of wine, and feeling happy or relaxed. Take a look at another image labeled in scared. Here, blue and bright colors consider the image as the one with the happy emotion, but the composition of the trees may make the image poster feel scared. In this sense, our model of focusing on image color and content through image objects shows that it can play some role in grasping posters' emotions.

5 DISCUSSION AND CONCLUSION

5.1 RQ1: Colors and content of SNS images

We examined the color distribution by illustrating the SNS images of each emotion in the color space. The results confirm that the color frequency of the SNS image of all emotions is dominated by black and white. Happy, sad, and scared show a wide distribution of warm reds and yellows and have a common emotion of calm. Conversely, angry, excited, and satisfied have a common emotion of exciting. This result is different from what we normally think of as positive (i.e., happy, excited, satisfied) and negative emotions (i.e., sad, angry, scared).

This result can be explained by the difference in the image characteristics between the images from SNS and those used in the prior studies (i.e., IAPS, ArtPhoto, and Abstract Painting datasets). Many times, SNS users post an image with their emotions and display an additional description thereof through a hash tag. A user's explicit emotion, expressed in such a hash tag may not always coincide with what viewers think. This can be explained as a paradoxical emotion expression that frequently occurs in an image-based SNS.

As a result of topic modeling, we confirmed that the topics viewed in positive and negative emotions are distinctly different. The topics that appeared mostly in positive emotions were person, fashion, and activity. In the content of each topic, person and act (e.g., man, woman, smiling, posing) appeared in person and clothes and appearance (e.g., accessory, dress) appeared in fashion, activity, travel, and sport. This illustrates that the SNS users communicate their positive emotions through the expression of the person, their appearance, and activities. On the other hand, the main topics appearing in negative emotions are abstract and stuff. Content from abstract and stuff mainly includes the description of situations (e.g, stationary, finned) and objects (e.g., clock, oddments, toy). This infers that SNS users use some objects as a means to express their negative emotions. Each emotion's topic showed different characteristics. Food shows a greater weight on satisfied, than other positive emotions. Person and outdoor show greater weights on scared than other negative emotions. Regarding content, even in the same category, cats and dogs appear more in positive emotions, while reptiles and lions appear more in negative emotions.

This suggests that SNS users tend to express each emotion in relation to a specific topic. In conclusion, the color and content of the SNS image is closely related to the user's emotions.

5.2 RQ2: Emotion classification

Given that the previous CNN-based emotion classification model has a true positive rate performance of 0.70 to 0.90 for each emotion, our emotion classification model illustrated comparable results with a performance of 0.77 to 0.85 (F1 score). Moreover, while the model performance has not been adequately verified due to the high imbalance in the sample size for each emotion in previous studies, our models are based on a sufficient and same amount of samples and on a 5 cross-validation. For this reason, our classification model is built and tested in a more comprehensive fashion.

The performance of our model, which was found to be greater than that of other CNN models, highlights that the color and content of the images are an important feature in emotion classifications. This confirms the possibility of emotion classification through the color and content of SNS images.

5.3 Study implications

Many prior studies have suggested emotion classification models through a small dataset with types of emotions that can be relatively easily identified. However, these datasets are likely to be of a different nature from the vast amounts of SNS images actually produced by the user, in terms of labeled datasets, from the viewpoint of the image viewer. Our emotion model reflects the viewpoint of the image poster, rather than that of the viewer. We believe the model could be well applied in various fields.

Advertising through SNS has become one of the most effective marketing tools. Hence, advertisers have tried to maximize the effects of advertisements, where SNS users' expose their emotions early on. Alternatively, recently popularized music and clip streaming services also provide services for recommending music or clips based on feelings. Therefore, it is important to grasp the emotions of the users for the emotion-based recommendation systems that recommend advertisements, music, and clips based on these emotions. In addition, mental care services, such as warnings of depression through human emotional changes, and other services (e.g., automatically adjusting the lighting of the interior space according to the user's emotions), in accordance with the recent development of IoT can be utilized in various ways. In order to grasp the emotions of these users, it is a good alternative to grasp the emotions through the image of the SNS space that users use. In this respect, our research can be applied widely.

6 CONCLUSIONS

Although our research has shown a number of insights on user's emotion recognition and modeling, there are some limitations that we plan to address in the future. Our models used the color and content features which could be improved by considering other features. For example, combinations of image objects, assigning weights to each object depending on its importance in the image, using other tags/captions attached to the same image, could be considered as new features. In addition, we collected images using one hashtag for each emotion. Since there are many other hashtags that refer to the same emotions (e.g., happy-delighted, sad-depressed), we plan to use it to collect additional data and build emotion classification models based on the more data. In summary, our research highlights the importance of considering the poster's emotions. This will better reflect the dynamics of SNS images regarding volume and diversity as well as suitably applied to many domains (e.g., emotion detection, recommendation).

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (NRF-2017M3C4A7069440, NRF-2017M3C4A7083529, No.2018R1A5A7059549).

REFERENCES

[1] Morgan Ames and Mor Naaman. 2007. Why We Tag: Motivations for Annotation in Mobile and Online Media. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 971–980.

[2] Barry J Babin, David M Hardesty, and Tracy A Suter. 2003. Color and Shopping Intentions: The Intervening Effect of Price Fairness and Perceived Affect. *Journal of business research* 56, 7 (2003), 541–551.

[3] Matteo Baldoni, Cristina Baroglio, Viviana Patti, and Paolo Rena. 2012. From Tags to Emotions: Ontology-driven Sentiment Analysis in The Social Semantic Web. *Intelligenza Artificiale* 6, 1 (2012), 41–54.

[4] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent Dirichlet Allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.

[5] Damian Borth, Tao Chen, Rongrong Ji, and Shih-Fu Chang. 2013. Sentibank: Large-scale Ontology and Classifiers for Detecting Sentiment and Emotions in Visual Content. In *Proceedings of the ACM International Conference on Multimedia*. 459–460.

[6] Damian Borth, Rongrong Ji, Tao Chen, Thomas Breuel, and Shih-Fu Chang. 2013. Large-scale Visual Sentiment Ontology and Detectors using Adjective Noun Pairs. In *Proceedings of the ACM International Conference on Multimedia*. 223–232.

[7] Ming Chen, Lu Zhang, and Jan P Allebach. 2015. Learning Deep Features for Image Emotion Classification. In *Proceedings of IEEE International Conference on Image Processing*. 4491–4495.

[8] François Chollet. 2016. Xception: Deep Learning with Depthwise Separable Convolutions. *arXiv preprint* (2016).

[9] Mihaly Csikszentmihalyi and Eugene Halton. 1981. *The Meaning of Things: Domestic Symbols and The Self.* Cambridge University Press.

[10] Paul Ekman. 1992. An Argument for Basic Emotions. *Cognition & emotion* 6, 3-4 (1992), 169–200.

[11] Richard Hanna, Andrew Rohm, and Victoria L Crittenden. 2011. We're All Connected: The Power of The Social Media Ecosystem. *Business horizons* 54, 3 (2011), 265–273.

[12] Simon Haykin and Neural Network. 2004. A Comprehensive Foundation. *Neural networks* 2, 2004 (2004), 41.

[13] Carroll E Izard. 1991. *The Psychology of Emotions.* Springer Science & Business Media.

[14] Carroll E Izard. 2007. Basic Emotions, Natural kinds, Emotion Schemas, and A New Paradigm. *Perspectives on psychological science* 2, 3 (2007), 260–280.

[15] Jin Yea Jang, Kyungsik Han, Patrick C Shih, and Dongwon Lee. 2015. Generation Like: Comparative Characteristics in Instagram. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 4039–4042.

[16] A. Kaplan and M. Haenlein. 2010. Users of The World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons* 53, 1 (2010), 59–68.

[17] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. 2014. Large-scale Video Classification with Convolutional Neural Networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 1725–1732.

[18] Diederik P Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* (2014).

[19] Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. 1997. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *NIMH Center for the Study of Emotion and Attention* (1997), 39–58.

[20] Xin Li, Lei Guo, and Yihong Eric Zhao. 2008. Tag-based Social Interest Discovery. In *Proceedings of the International Conference on World Wide Web*. 675–684.

[21] Xin Lu, Reginald B Adams, Jia Li, Michelle G Newman, and James Z Wang. 2017. An Investigation into Three Visual Characteristics of Complex Scenes that Evoke Human Emotion. In *Proceedings of the IEEE International Conference on Affective Computing and Intelligent Interaction*. 440–447.

[22] Xin Lu, Poonam Suryanarayan, Reginald B Adams Jr, Jia Li, Michelle G Newman, and James Z Wang. 2012. On Shape and The Computability of Emotions. In *Proceedings of the ACM International Conference on Multimedia*. 229–238.

[23] Jana Machajdik and Allan Hanbury. 2010. Affective Image Classification using Features Inspired by Psychology and Art Theory. In *Proceedings of the ACM International Conference on Multimedia*. 83–92.

[24] Aqdas Malik, Amandeep Dhir, and Marko Nieminen. 2016. Uses and Gratifications of Digital Photo Sharing on Facebook. *Telematics and Informatics* 33, 1 (2016), 129–138.

[25] W Glynn Mangold and David J Faulds. 2009. Social Media: The New Hybrid Element of The Promotion Mix. *Business horizons* 52, 4 (2009), 357–365.

[26] KAYA NAz and Helena Epps. 2004. Relationship between Color and Emotion: A Study of College Students. *College Student J* 38, 3 (2004), 396.

[27] Oded Nov, Mor Naaman, and Chen Ye. 2008. What Drives Content Tagging: The Case of Photos on Flickr. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 1097–1100.

[28] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-scale Image Recognition. *arXiv preprint arXiv:1409.1556* (2014).

[29] Aaron Smith, Monica Anderson, and Tom Caiazza. 2018. Social Media Use in 2018.

[30] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2818–2826.

[31] Joshua B Tenenbaum, Vin De Silva, and John C Langford. 2000. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *science* 290, 5500 (2000), 2319–2323.

[32] Kate Vaisutis, Margot Brereton, Toni Robertson, Frank Vetere, Jeannette Durick, Bjorn Nansen, and Laurie Buys. 2014. Invisible Connections: Investigating Older People's Emotions and Social Relations around Objects. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1937–1940.

[33] Patricia Valdez and Albert Mehrabian. 1994. Effects of Color on Emotions. *Journal of experimental psychology: General* 123, 4 (1994), 394.

[34] José Van Dijck. 2013. *The Culture of Connectivity: A Critical History of Social Media.* Oxford University Press.

[35] Anita Whiting and David Williams. 2013. Why People Use Social Media: A Uses and Gratifications Approach. *Qualitative Market Research: An International Journal* 16, 4 (2013), 362–369.

[36] Can Xu, Suleyman Cetintas, Kuang-Chih Lee, and Li-Jia Li. 2014. Visual Sentiment Prediction with Deep Convolutional Neural Networks. *arXiv preprint arXiv:1411.5731* (2014).

[37] Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang. 2016. Building a Large Scale Dataset for Image Emotion Recognition: The Fine Print and The Benchmark.. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 308–314.

[38] Sicheng Zhao, Yue Gao, Xiaolei Jiang, Hongxun Yao, Tat-Seng Chua, and Xiaoshuai Sun. 2014. Exploring Principles-of-art Features for Image Emotion Recognition. In *Proceedings of the ACM International Conference on Multimedia*. 47–56.